Galectin-3-Binding Glycomimetics that Strongly Reduce Bleomycin-Induced Lung Fibrosis and Modulate Intracellular Glycan Recognition

A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Galectin-3-binding glycomimetics that strongly reduce bleomycin-induced lung fibrosis and modulate intracellular glycan recognition

Authors: Tamara Delaine; Patric Collins; Alison MacKinnon; G Sharma; John Stegmayr; Vishal K Rajput; Santanu Mandal; Ian Cumpstey; Amaia Larumbe; Bader A Salameh; Barbro Kahl-Knutsson; Hilde Van Hattum; Monique Van Scherpenzeel; Roland J Pieters; Tariq Sethi; Hans Schambye; Stina Oredsson; Hakon Leffler; Helen Blanchard; Ulf Nilsson

To be cited as: ChemBioChem 10.1002/cbic.201600285

Link to VoR: http://dx.doi.org/10.1002/cbic.201600285

A Journal of

ChemPubSoc Europe

www.chembiochem.org

WILEY-VCH

Galectin-3-binding glycomimetics that strongly reduce bleomycin-induced lung fibrosis and modulate intracellular glycan recognition

T. Delaine,[a] P. Collins,[b] A. MacKinnon,[c] G. Sharma,[d] J. Stegmayr,[d] V. K. Rajput,[a] S. Mandal,[a] I. Cumpstey,[a] A. Larumbe,[a] B. A. Salameh,[a]† B. Kahl-Knutsson,[d] H. van Hattum,[e] M. v. Scherpenzeel,[e]‡ R. J. Pieters,[e] T. Sethi,[f] H. Schambye,[g] S. Oredsson,[h] H. Leffler,[d] H. Blanchard,*[b] and U. J. Nilsson*[a]

Abstract: Discovery of glycan-competitive galectin-3-binding compounds that attenuate lung fibrosis in a murine model and that block intracellular galectin-3 accumulation at damaged vesicles, hence revealing galectin-3-glycan interactions being involved in fibrosis progression and in intracellular galectin-3 activities is reported. Sixteen 3,3´-bis-(4-aryl-triazol-1-yl)-thiodigalactosides were synthesized and evaluated as antagonists of galectin-1, 2, 3, 4 N-terminal, 4 C-terminal, 7, 8 N-terminal, 9 N-terminal, and 9 C-terminal domains. Compounds displaying low-nM affinities for galectin-1 and 3 were identified in a competitive fluorescence anisotropy assay. X-ray structural analysis of selected compounds in complex with galectin-3 and galectin-3 mutant binding experiments revealed that both aryl-triazolyl moieties and fluoro-substituents of the compounds are involved in key interactions responsible for the exceptional affinities for galectin-3. The most potent galectin-3 antagonist was demonstrated to act in an assay monitoring galectin-3 accumulation upon amitriptyline-induced vesicle damage, visualizing a biochemical/medical relevant intracellular lectin-carbohydrate binding event and that it can be blocked by a small molecule. The same antagonist administered intratracheally attenuated bleomycin-induced pulmonary fibrosis in a mouse model with a dose-response profile comparing favorably to orally administration of the marketed anti-fibrotic compound pirfenidone.

[a] Dr. T. Delaine, Dr. V.K. Rajput, Dr. S. Mandal, Dr. I. Cumpstey, Dr. A. Larumbe, Dr. B.A. Salameh, Prof. Dr. U.J. Nilsson
Centre for Analysis and Synthesis, Department of Chemistry
Lund University
POB 124, SE-221 00 Lund, Sweden.
E-mail: ulf.nilsson@chem.lu.se
[b] Dr. P. Collins, Dr.H. Blanchard
Institute for Glycomics
Griffith University, Gold Coast Campus
Queensland 4222, Australia
E-mail: h.blanchard@griffith.edu.au
[c] Dr. A. MacKinnon
MRC Centre for Inflammation Research, The Queen's Medical Research Institute
University of Edinburgh
Edinburgh, UK
[d] Dr. G. Sharma, J. Stegmayr, Mrs. B. Kahl-Knutsson, Prof. Dr. H. Leffler
Department of Laboratory Medicine, Section MIG
Lund University
BMC-C1228b, Klinikgatan 28, SE-221 84 Lund, Sweden
[e] Dr. H. van Hattum, Dr. M. van Scherpenzeel, Prof. Dr. R.J. Pieters
Department of Medicinal Chemistry and Chemical Biology, Utrecht Institute for Pharmaceutical Sciences
Utrecht University
P.O. Box 80082, 3508 TB Utrecht, The Netherlands
[f] Prof. Dr. T. Sethi
Department of Respiratory Medicine and Allergy
Kings College, Denmark Hill Campus
London, UK
[g] Dr. H. Schambye
Galecto Biotech ApS
COBIS
Ole Maaloes vej 3, Copenhagen N, DK-2200, Denmark
[h] Prof. Dr. S. Oredsson
Department of Biology
Lund University
POB 118, Lund, 221 00 Sweden
† Present address: Chemistry Department, The Hashemite University, PO Box 150459, Zarka 13115 Jordan
‡ Present address: Translational Metabolic Laboratory, Radboud University Medical Center, Nijmegen, The Netherlands

Supporting information for this article is given via a link at the end of the document.

Introduction

The galectins are a family of proteins that have the ability to cross-link β-D-galactopyranoside-containing glycoproteins (and other glycoconjugates) to form lattices[1] and thereby modulate glycoprotein localization, transport, and residence times in cellular compartments and at surfaces.[2] Cross-linking of glycoproteins by galectins can occur due to the galectins' capability to present multiple carbohydrate recognition sites (CRD) depending on their type. Prototype galectins (1, 2, 7, 10, 11, 13, 14, and 15) contain one CRD but dimerize depending on their concentration and ligand density. The tandem-repeat galectins (4, 5, 6, 8, 9, and 12) contain and present two CRD's, and the chimera-type galectin-3 CRD is linked to a glycine/proline-rich collagen-like N-terminal domain that enables oligomerization.

This organizational lattice-forming role of the galectins influences glycoprotein activities and the duration thereof, as well as glycoprotein intracellular trafficking and sorting. This manifests itself in different effects on the cellular level that depend on a match between galectin type and expression, as well as on the glycan structures in the cell. For example, galectin-glycoconjugate interactions control cell properties and functions, cell adhesion, have immunomodulatory effects[3] and effects on tumor growth and metastases.[4] The cellular mechanisms and roles in inflammation and cancer point to the

use of galectin CRD antagonists as therapeutic agents and several *ex vivo*[5] and *in vivo*[6] studies of the most studied and well characterized galectin-3 have corroborated such hypotheses.

Among attempts to develop small and drug-like molecules as galectin-3 antagonists, substitution of galactose[6a, 7] (as such, or part of lactose or *N*-acetyl-lactosamine (LacNAc)) and 3,3-di-substitution of thiodigalactoside[8] have proven to be successful. In particular, high-affinity small-molecule galectin-3 antagonists with sub-μM affinities have been discovered by appending aromatic amido groups or 4-amido-1,2,3-triazolyl groups at both C-3 carbons in thiodigalactoside.[8a, 8b, 8d] Here, we present 4-aryl-1,2,3-triazolyl thiodigalactoside-based derivatives as significantly improved antagonists with selectivity for galectin-1 and 3. Furthermore, an investigation based on three X-ray structures of galectin-3 in complex with inhibitors and on galectin-3 mutant studies revealed that the aryl-triazolyl groups form affinity-enhancing interactions with arginine side-chains and with ß-strand backbones. One selected compound was demonstrated to function intracellularly in an amitriptyline-induced vesicle damage assay and to reduce fibrosis levels in a murine bleomycin lung fibrosis model.

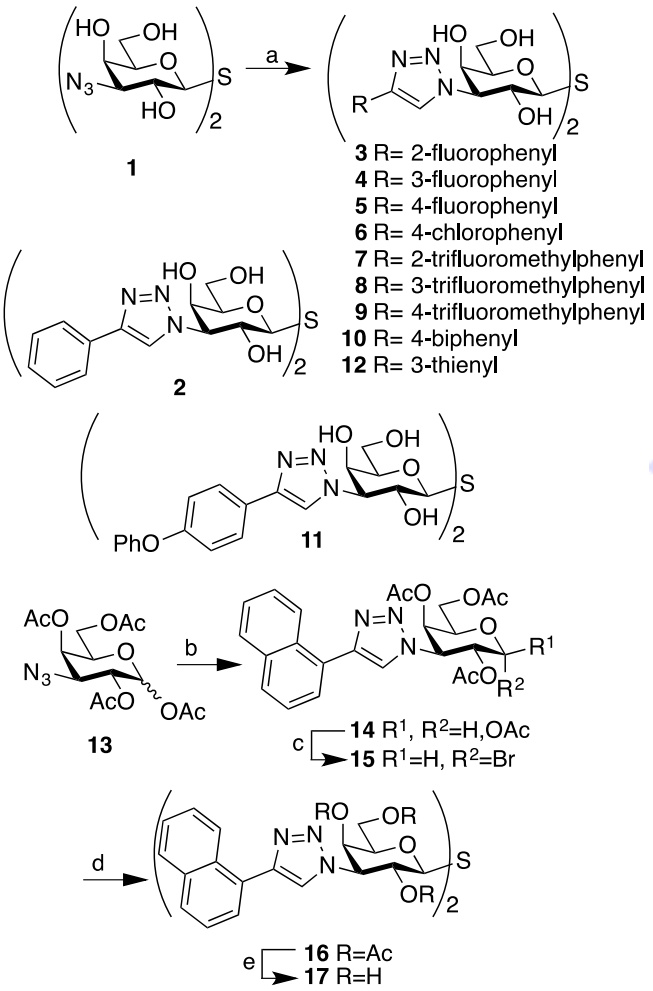
Results and Discussion

Synthesis

The ditriazolyl-thiodigalactosides **3**-**10** and **12** were synthesized by Cu(I)-catalyzed cycloadditions between the known diazide **1**[9] and phenylacetylenes (Scheme 1), while synthesis of the unsubstituted phenyltriazole **2** and the phenoxyphenyltriazole **11** have been reported earlier[10] (Table 1). The 1-naphthyl-triazole **17** was synthesized essentially following a previously published alternative procedure[8d] that involved cycloaddition of 1-ethynylnaphthalene with the acetylated galacto azide **13**[11] to give the triazole **14**. Bromination of **14** and subsequent double substitution of the bromide **15** with sodium sulfide resulted in the thiodigalactoside **16** in a moderate yield. De-*O*-acetylation of **16** gave the target 1-naphthyl-triazole **17**.

Galectin affinities and structure-activity relationships

With a panel of bis-aryltriazolyl thiodigalactosides **2**-**12** and **17** at hand, affinities towards galectin-1, 2, 3, 4 N- and C-terminal domains, 7, 8 N-terminal domain, and 9 N- and C-terminal domains were determined in a competitive protein-binding assay based on fluorescence anisotropy as earlier described in detail.[12] Except for galectin-8N, all investigated galectins bound all, or most, of the inhibitors **2**-**12** and **17** with affinities significantly greater than those of the parent unsubstituted thiodigalactoside (Table 1). Galectin-1 bound all phenyl-triazoles, unsubstituted or with smaller substituents, (**2**-**9**) with indeed high affinities, while larger substituents (**10** and **11**) significantly reduced affinity. Interestingly, 3- and 4-fluorinated phenyl compounds **4** and **5** turned out to be the only ones better than the unsubstituted phenyl **2** and the 2-fluorophenyl derivative **3**, with dissociation constants as low as 12 nM (3-fluorophenyl **4**).



Scheme 1. Reagents and conditions: (a) Alkyne, CuI, Et_3N, DMF; (b) Alkyne, CuI, DIPEA, toluene, 65-80°C; (c) HBr/AcOH; (d) Na_2S, MS 4Å, MeCN; (e) $BuNH_2$, MeOH.

This suggests that one or both of the galectin-1 subsites that accommodate the phenyl groups of **2**-**11** are tight with limited possibilities for substitutions as suggested based on earlier analyses of **2** and **11**.[10] The preference for substitution position of the fluorophenyl derivatives **3**-**5** (m<p<o) is reflected in the corresponding trifluoromethyl series **7**-**9**, albeit at somewhat higher K_d values. Although the fluorophenyl-carrying **4** and **5** indeed reach low nM affinities for galectin-1, more noteworthy is the even higher affinity of the thienyl compound **12**. This compound provided near quantitative inhibition at all concentrations tested and an accurate dissociation constant could not be reliably calculated. Hence, the dissociation constant could only be estimated to be less than 10 nM, which is at least 2400–fold better than the reference unsubstituted thiodigalactoside and natural disaccharide ligands.

Table 1. Dissociation constants (μM) for **2**-**12**, and **17** and galectin-1, 2, 3, 4, 4 N-terminal domain, 4 C-terminal domain, 7, 8 N-terminal domain, and 9 N- and C-terminal domain determined with a competitive fluorescence anisotropy assay.[12]

	Galectin								
	1	2	3	4N	4C	7	8N	9N	9C
2	0.049[10]	1.2±0.27	0.044[10]	2.6±0.25	0.19±0.015	1.2±0.21	>100	2.3±0.23	0.98±0.16
3	0.31±0.045	0.98±0.12	0.19±0.033	2.9±0.63	0.39±0.083	1.6±0.17	83±9.9	1.0±0.28	1.4±0.22
4	0.012±0.003	>5	0.014±0.003	0.17±0.029	0.14±0.042	1.9±0.38	86±8.8	0.68±0.34	0.12±0.015
5	0.027±0.003	>5	0.034±0.0058	0.65±0.24	0.073±0.003	4.2±0.79	104±15	1.8±0.20	0.58±0.10
6	0.33±0.030	10±3.9	0.19±0.044	7.2±1.2	4.7±0.34	3.4±0.56	210±19	2.8±0.76	1.8±0.36
7	1.4±0.0.70	1.5±0.44	0.38±0.039	3.8±0.35	4.9±0.87	2.0±0.22	200±11	0.69±0.017	0.78±0.20
8	0.45±0.039	1.6±0.24	0.23±0.036	15±3.0	1.5±0.27	7.8±1.9	>500	3.7±1.0	4.2±1.2
9	1.2±0.19	1.4±0.29	0.25±0.043	7.9±2.5	8.6±1.5	9.1±3.6	>500	11±0.64	1.8±0.14
10	110±17	>500	770±8.4	>1000	650±39	>1000	>10000	>1000	>1000
11	84[10]	32±9.5	0.36[10]	>500	>500	>500	440±14	>500	240±11
12	<0.010	>5	0.065±0.09	3.8±0.57	0.19±0.035	2.9±0.49	120±6.2	2.2±0.20	0.46±0.037
17	nd[a]	nd	0.98±0.023	nd	nd	nd	120±1.8	0.31±0.006	nd
TDG[b]	24[8a]	340±19	49[8a]	410±21	980±70	160[8a]	61[8a]	38[8a]	42±1.1

[a] Not determined. [b] Thiodigalactoside.

In stark contrast, galectin-2 was inhibited with only micromolar K_d values by any of **2**-**12** and **17**. The best compound, the 2-fluorophenyl compound **3**, reached only a moderate affinity of about 1 μM, which is nevertheless significantly better than the parent unsubstituted thiodigalactoside reflecting the presence of positive interactions between **2**-**17** and this galectin. Galectin-3 was well inhibited by several compounds and interestingly showed a selectivity profile similar to galectin-1; phenyl moieties carrying small substituents (**2**-**9**), as well as the thienyl moiety (**12**), conferred high affinity, while phenyls carrying larger substituents (**10**-**11**), as well as the naphthyl (**17**), were less efficiently bound by this galectin. A notable difference is, however, that while the biphenyl **10** is virtually detrimental to binding (as for galectin-1), the 4-phenoxy-substituted phenyl **11** is reasonably well tolerated by galectin-3, with a sub-μM affinity, which is not the case for galectin-1. Hence, compound **11** displays, as earlier reported,[10] an important more than 50-fold selectivity for galectin-3 over galectin-1. The reverse situations holds for the thienyl **12**, which inhibits galectin-3, albeit with an affinity of 65 nM, but still less well than galectin-1. Hence, the thienyl derivative **12** has a clear selectivity for galectin-1 over galectin-3 and thus through proper choice of aryl substituents on the triazole rings, selectivity for galectin-1 (by **12**) or galectin-3 (by **11**) is achieved.

Both CRD of the tandem-repeat galectin-4 were evaluated and the N-terminal domain did recognize compounds **2**-**12** and **17** with moderate affinities in the low-medium μM range, which for all compounds is better than the parent thiodigalactoside. Reflecting the difference in fine-specificity between the two galectin-4 domains, the C-terminal domain revealed mid-nM affinities for several compounds. As observed for galectin-1 and 3, aryltriazoles carrying no or smaller substitutents at the aryl moiety (**2**-**5** and **12**) were identified as the best inhibitors. Again, this suggests that one or both of the aryl-accommodating sites of galectin-4C can harbor only smaller structures. The 4-fluorophenyl derivative **5** stands out as a most potent galectin-4C with a K_d of 73 nM, which suggests a specific fluorine interaction and/or an ideal steric fit by the 4-fluoro substituent. Most likely, efficient inhibition of one domain[13] will be sufficient to block physiological/biological effects by galectin-4. Galectin-7 binding is enhanced by the 4-aryl-triazolyl groups of **2**-**9**, **12**, and **17**, while the sterically more demanding compounds **10**-**11** are virtually non-binding. Similar observations were made for galectin-9N and 9C, which both bind several inhibitors with sub to low μM affinities. In contrast to galectin-4, no clear selectivity between the two domains of galectin-9 was observed.

Overall, the 4-aryl-triazolyl thiodigalactosides **2**-**12** and **17** delivered inhibitors significantly more potent than thiodigalactoside itself against galectin-1, 2, 3, 4N, 4C, 7, 9N, and 9C and more potent than the corresponding galactoside monosaccharide derivative against galectin-3, 7, and 9N (*c.f. e.g.* the monosaccharide corresponding to **2** show K_d of 150, 1700, and 1300 μM, respectively, against these four galectins[7a,8d]). In particular, galectin-1 and 3 were well inhibited with several compounds showing low nM affinity. The inhibition potency against galectin-3 even surpassed our earlier described corresponding 4-amido-triazolyl-[8d] and 4-aryl-triazolyl-thiodigalactoside[10] derivatives. Several compounds indeed possess a clear selectivity for these two galectins, while the

selectivity between them is limited except for the thienyl derivative **12** and the phenoxy derivative **11** that displayed moderate selectivity for galectin-1 and 3, respectively.

Structural studies

The X-ray structures were determined of three selected high-affinity compounds, the 3- and 4-fluophenyl derivatives **4** and **5** and the thienyl derivative **12**, in complex with galectin-3. Initial refinements of X-ray diffraction data (1.5–1.6 Å resolution, Supplementary Table S1) produced clear difference electron density within the galectin binding sites revealing the bound ligands **4**, **5**, and **12** (Figure 1). In all cases, electron density is clearly evident for the thiodigalactoside core of all three ligand **4**, **5**, and **12** (except for the solvent orientated C6 hydroxyl), and for both triazole rings. The thiodigalactoside core of each ligand **4**, **5**, and **12** is in an identical binding mode to that observed in the previously reported galectin-3-thiodigalactoside complex[14] and forms identical protein–ligand interactions, confirming that they do not act as divalent ligands. Electron density is also defined for the aromatic rings extending from one of the triazole C4 atoms of **4**, **5**, and **12** towards the Arg144 side chain. The second thiophene or fluorophenyl rings of the ligands **4**, **5**, and **12**, positioned above the salt bridge between Glu165 and Arg186, are less clearly defined by the initial difference electron density maps compared to the rest of the ligand (upper region of the ligands in Figure 1a–c). The position of the thiophene ring of **12** and the 3-fluorophenyl ring of **4** is evident in the region near Glu165 and Arg186 in difference electron density maps when scaled to 2.5 σ (calculated prior to addition of the ligand to the model) and refinements with the ligand included in the model show the thiophene and 3-fluorophenyl rings defined by $2mF_o-DF_c$ electron density when scaled at 0.7 σ. Additional weak $2mF_o-DF_c$ electron density appears near Glu165 and Arg186 of **4** after refinement, indicating a possible alternate conformation for the ring, however, the electron density is not clear enough to confidently model two alternate conformations for the ligand.

The 4-fluorophenyl ring of **5** near the Glu165–Arg186 salt bridge is poorly defined by the initial difference electron density. Refinement of the model with the ligand **5** in place, but excluding the 4-fluorophenyl ring near Glu165–Arg186, results in additional difference electron density that indicates the general location of the ring, and refinements with the ring included in the model resulted in weak $2mF_o-DF_c$ electron density (0.7 σ) that supports the location of the ring. However, there is clearly a higher degree of disorder for this part of the ligand. This may initially appear counter-intuitive as known ligands with aromatic groups near the Glu165–Arg186 salt bridge region of galectin-3 have shown enhanced affinities (for example the diamido-thiodigalactosides[8a, 8b] and aromatic lactose 2-*O*-esters[15]). However, the interaction involves face-to-face stacking between the aryl-triazoles onto an extended surface of the π-system of the Glu165–Arg186 ion-pair, which could allow for the aryl-triazole to position over different segments of the large π-system with retained interaction free energies.

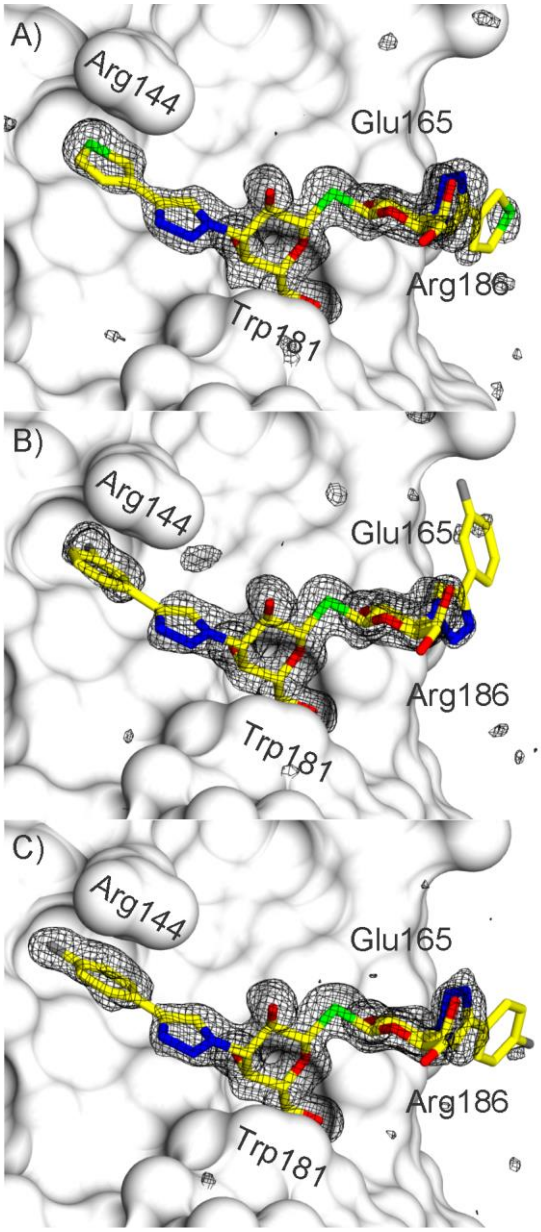


Figure 1. Difference electron density within the galectin-3 CRD binding sites showing bound A) **12**, B) **4**, and C) **5**. Difference electron density calculated from refinement with the ligand (stick representation) omitted from the model ($|Fo|-|Fc|$ α_{calc}; grey mesh, contoured at 3□) and the protein represented by a grey solvent-accessible surface.

The triazole ring of compounds **4**, **5**, and **12** located above His158 is orientated with the nitrogens positioned towards Trp181, which allows for formation of a water-mediated hydrogen bond between the N2 of the triazole and the nitrogen atom of Trp181 (Figure 2 a-c). The triazole ring of compounds **4**, **5**, and **12** located near the in-plane Glu165-Arg186 salt bridge shows two alternate conformations that stack face-to-face to the in-plane Glu165-Arg186 salt bridge. In the **12** and **5** complexes, the triazole nitrogens are directed towards Glu165, while in the

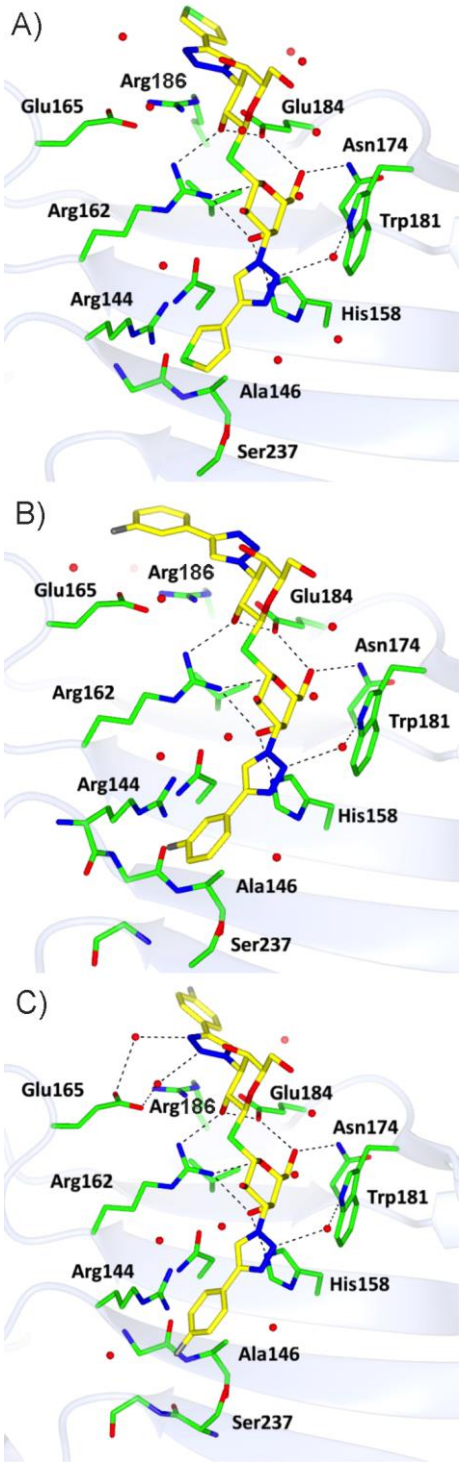


Figure 2. Galectin-3 CRD binding site interactions with A) **12**, B) **4**, and C) **5**. H-bond interactions between ligand (yellow bonds) and protein/water (green bonds) are shown as dashed lines.

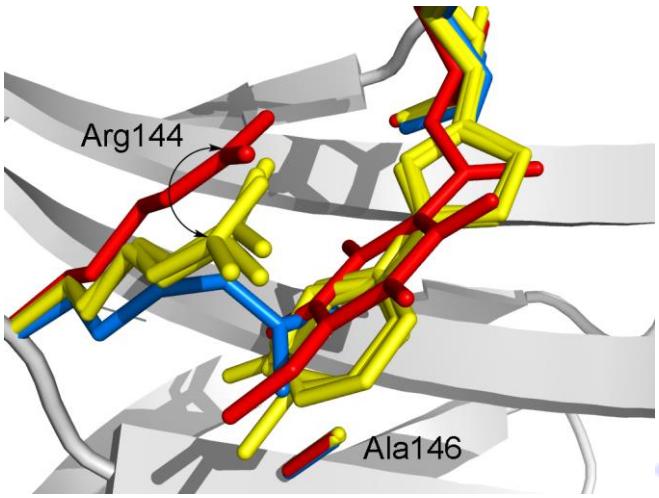


Figure 3. Superimposed view of the galectin-3 CRD binding site in the region of Arg144 for the complex with **4**, **5**, and **12** (yellow), the 3'-(2,3,5,6-tetra-fluoro-4-methoxybenzamido) LacNAc derivative (red, PDB ID: 1KJR), and lactose (blue, PDB ID: 3ZSJ).

complex with **4** the ring is flipped with the triazole nitrogens close to Arg186. The orientation of one of the triazole rings in **12** and **5** results in contacts with both Glu184 and Arg186, while in **4** the triazole is in contact with Arg186 only.

In all three galectin-3 complexes with **4**, **5**, and **12**, the ligands induces a conformational change in Arg144 (Figure 3) similar to that reported for the galectin-3 CRD structure in complex with 3'-(2,3,5,6-tetra-fluoro-4-methoxybenzamido)-LacNAc derivative (PDB ID: 1KJR).[7c] One of the terminal aromatic rings of the ligands **4**, **5**, and **12** fits into a pocket that is exposed by the Arg144 conformational change and move away from the protein surface and forms a face-to-face stacking interaction with Arg144 in a similar manner as observed for the corresponding complex with 3'-(2,3,5,6-tetra-fluoro-4-methoxybenzamido)-LacNAc derivative. However, in the structures of the complexes with **4**, **5**, and **12**, the additional length granted by the triazole linker allows the terminal aromatic rings to extend deeper into the pocket exposed by the Arg144 move, which allows for the formation of an additional contact with Ala146 that is not observed in the 3'-(2,3,5,6-tetra-fluoro-4-methoxybenzamido)-LacNAc complex. Additionally, although the conformational change of Arg144 in the complexes with **4**, **5**, and **12** is overall similar to that earlier observed for 3'-(2,3,5,6-tetra-fluoro-4-methoxybenzamido)-LacNAc derivative,[7c] small differences are apparent. The Arg144 has moved in the complexes with **4**, **5**, and **12** compared to the complex with the 3'-(2,3,5,6-tetra-fluoro-4-methoxybenzamido)-LacNAc derivative (1.5–2.0 Å ζ-carbon to ζ-carbon distance) such that the guanidino group maintains its position directly above the aromatic ring of the ligand (Figure 3). One thiophene ring of **12** is orientated to deeply bury the sulfur atom in the pocket exposed by the Arg144 move, as is the fluorine atom in the 3-fluorophenyl ring of **4**. The SAD LLG map calculated for the **12** complex confirms the orientation of the thiophene ring showing a clear peak positioned at the location of the sulfur atom within the pocket near Arg144. The fluorine of **4**

below Arg144 is situated at a distance of 3.9Å and 3.4Å and at angles of 155° and 147° from the backbone carbonyls of Arg144 and Ile145, respectively, which suggests the formation of two orthogonal multipolar interactions.[16] The fluorine atom of the 4-fluorophenyl ring in **5** is directed towards Gly238 and Ser237 and makes contact with the *α*-carbon of Gly238 and is also positioned well for forming an orthogonal dipolar interaction with the Ser237 carbonyl (distance 3.6Å and angle 140°). Furthermore, the guanidinium ion of arginine side chains has been proposed to be highly fluorophilic, as fluorine atoms of fluorinated pharmaceuticals have been observed to be close to guanidinium moieties in proteins.[16a, 17] Finally, fluorination typically results in increased lipophilicity[16a, 17] and fluorinated hydrocarbons are in general poorly solvated in water,[18] which would support a conclusion that burying fluorinated lipophilic ligand parts is important for achieving high affinity of **4** and **5** for galectin-3. The equivalent of Arg144 is absent in some galectins[19] and consequently targeting ligand interactions to this region and engaging Arg144 through cation-π interactions is proposed as a means of enhancing galectin binding selectivity.

Galectin-3 mutant studies

The X-ray structures of galectin-3 revealed that the aryl-triazoles of **4**, **5**, and **12** stacked face-to-face onto two (Arg144 and Arg186) arginine guanidinium groups. In case of galectin-3, the two 3-fluorophenyl moieties of **4** have different stacking modes with the two Arg144 and Arg186: One 3-fluorophenyl moiety is stacked on top of Arg186 guanidinium group, while the other 3-fluorophenyl moiety is inserted between the protein surface (backbone) and Arg144 guanidinium group (Figure 1b, 2b, and 3). In order to obtain further understanding about the nature of the aryl-triazole arginine stacking interactions, we determined the affinity of **4** for four galectin-3 mutants, R144K, R144S, R186K, and R186S (Table 2). The R144S and R186S mutants were chosen because the side-chain is removed without introducing a very non-polar surface and the R144K and R186K mutants were chosen because the cationic nature of the side-chain is retained while the planar π-system of the guanidino group is removed. The effect of the R144S mutant is minimal, which suggests that the stacking of Arg144 onto the 3-fluorophenyl group of **4** does not contribute significantly to the free energy of binding, while the surface complementarity and interactions of the 3-fluorophenyl group with the rest of the protein surface remains essentially unchanged.

Table 2. Dissociation constants (μM) for **4** against galectin-3 mutants determined with a competitive fluorescence anisotropy assay.[12]

wt	R144K	R144S	R186K	R186S
0.014±0.003	0.041±0.0045	0.017±0.0032	0.54±0.039	1.0±0.12

The R144K mutant binds **4** only about 3-fold less well than wild-type, which suggests that the lysine side-chain can, as an arginine side-chain, form cation-π interactions. However, in contrast to the arginine guanidine group, the lysine amino group obviously lacks a π-system and π-stacking capability may be the reason that the interaction with the 3-fluorophenyl group of **4** is possibly slightly less productive. The R186S mutant shows a large drop in affinity for **4**, clearly revealing that a 3-fluorophenyl stacking interaction onto Arg186 is an important contributor to the high affinity of **4** for galectin-3. The Arg186 side-chain guanidinim ion is, in contrast to the Arg144 side-chain, involved in an extensive network of in-plane bi-furcated ion-pairs (Figure 2), which form an extended π-system surface onto which a 3-fluorophenyl stacks in analogy with e.g. the acetamido group of *N*-acetyl-lactosamine[7c] and aromatic rings of 2-*O*-benzoyl lactose derivatives.[15] In the mutant R186S this extended π-system of bi-furcated ion-pairs is interrupted and the 3-fluorophenyl cannot form a beneficial stacking interaction. Instead, a poorly solvated cavity with poor complementarity to the 3-fluorophenyl group of **4** is present. Some binding affinity is regained in the R186K mutant as compared to the R186S mutant, which is presumably due the capability of the lysine side-chain to at least partly substitute and stabilize the Arg186 side-chain's key π-system-forming ion-pairing with the two surrounding Glu165 and Glu184 residues, as well as providing similar surface complementarity to the 3-fluorophenyl group of **4**.

In short, the high affinity of the 4-aryl-triazolyl thiodigalactosides, such as **4**, **5**, and **12**, for galectin-3 can be hypothesized, according to X-ray structural analysis of galectin-3 complexes and galectin-3 mutant studies, originating from several factors. First, ideal surface complementarity between the proteins and ligands (Figure 1) are, not unexpectedly, critical as this maximizes dispersion forces and presumably also beneficial desolvation effects. Stacking between galectin arginine side chain guanidinium functionalities and ligand phenyl-triazole moieties are probably important, as are fluorine orthogonal dipolar interactions[16c] with backbone carbonyls. Hence, while the core thiodigalactoside disaccharide mimics natural disaccharide ligand fragments (*e.g.* lactose and LacNAc) in terms of affinity contributions and structure, the appended non-carbohydrate aryl-triazole moieties engage in galectin-ligand interactions not seen in natural lectin-ligand complexes (*i.e.* predominantly hydrogen bonding and CH-π interactions), resulting in drastic affinity enhancements and enhanced selectivities.

Having discovered low-nM galectin antagonists, an important question of their efficiency for antagonizing galectin-glycoprotein interactions in biological systems was addressed with compound **4** in two models. First, an *in vitro* cell assay was developed with the goal of gaining new knowledge about galectin-3 putative intracellular glycan-binding activities and possible effects in cells challenged with vesicle-damaging agents. Second, to achieve further understanding of, as well as quantifying, the effects of antagonizing galectin-3 in an *in vivo* mouse model of bleomycin-induced idiopathic pulmonary fibrosis.[6b]

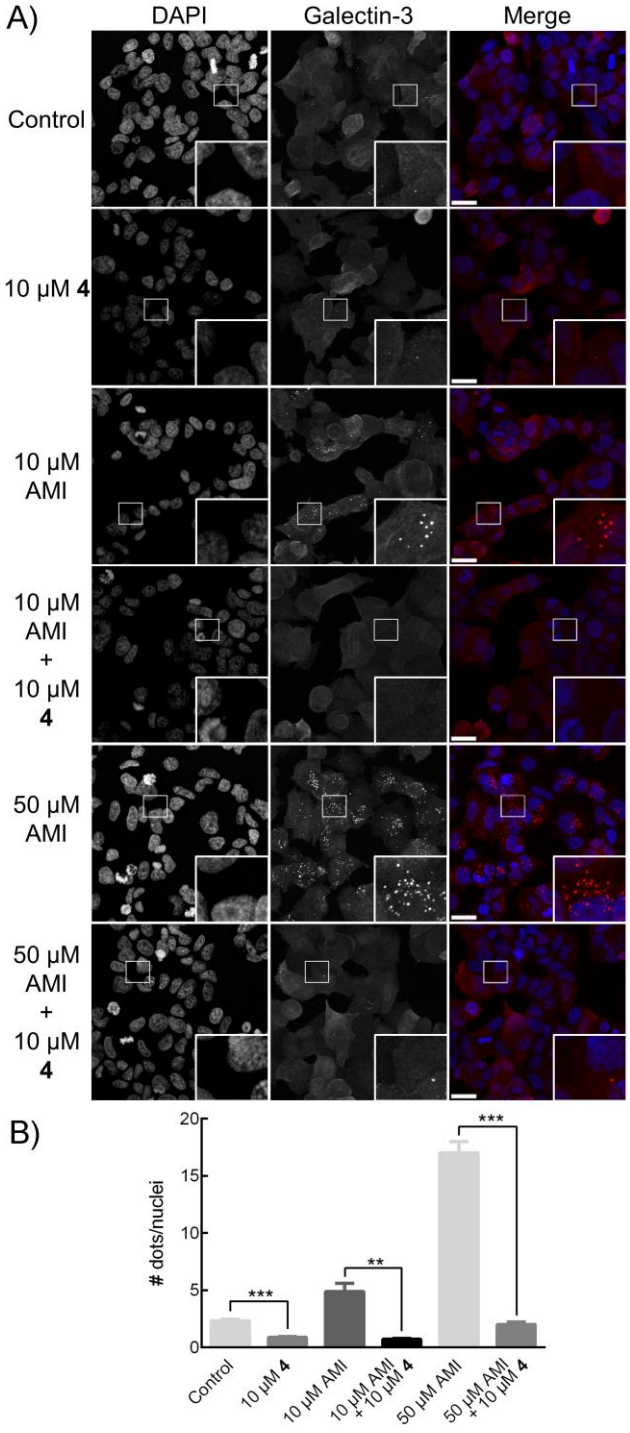


Figure 4. Inhibition galectin-3 accumulation around amitriptyline (AMI)-damaged vesicles in MCF-7 cells. Cells were treated with combinations of 10 μM compound **4** and 10 or 50 μM amitriptyline for 24 hours, control cells were treated with 0.1% v/v DMSO. A) Galectin-3 staining was visualized with anti-rat Alexa Fluor® 594 (red), whereas Hoechst (blue) was used to stain the nuclei. The immunofluorescence pictures displayed are representative for each treatment. Scale bars are equivalent to 20 μm. Small square inserts show which areas is magnified in each large square insert. B) The number of galectin-3 dots were counted manually using ImageJ in four different images for each experimental condition, and given as mean ± SEM. Each data set represents ~250 cells. **P<0.01, ***P < 0.001, Student's t-test.

Intracellular inhibition by galectin-3 antagonist in an amitriptyline-induced vesicle damage assay

Galectin accumulation around damaged vesicles in response to challenge by bacteria or chemical agents has been demonstrated in several studies and the formation of galectin-3[20] or galectin-8[21] puncta have been proposed as a novel marker for vesicular insult, regardless of the insult being of a bacterial[20b, 21] or chemical origin[20a]. Galectin-3 accumulation around damaged vesicles has, in addition, been shown to depend on glycan-binding, either by knock-down of certain glycosyltransferases[20b] or via knock-in of a galectin-3 mutant (R186S) with severely reduced affinity for endogenous glycans[20a]. Antagonizing effects on such glycan binding-dependent galectin-3 events on damaged vesicles[20b] could provide qualitative information on intracellular availability and activity of antagonists, such as compound **4**. Cationic amphiphilic drugs, including the tricyclic antidepressant amitriptyline, induce phospholipidosis and are speculated to accumulate in acidic lysosomes, and induce vesicle damage in tumor cell lines.[20a, 22] We postulated that treatment of cells with amitriptyline would induce formation of galectin-3 puncta in a similar fashion as other vesicular damaging agents, such as glycyl-l-phenylalanine 2-naphthylamide[21] and l-leucyl-l-leucine methyl ester[23]. Amitriptyline has the advantage of being more stable under the experimental conditions used and does not degrade in solution as *e.g.* glycyl-l-phenylalanine 2-naphthylamide. Furthermore, amitriptyline does not require the use of DMSO as co-solvent for solubilize the more commonly used peptidic vesicular damaging agents. Indeed, treating breast carcinoma MCF-7 cells with amitriptyline resulted in distinct accumulation of galectin-3 into vesicle-associated puncta, hypothetically within galectin-3:glycoprotein lattices, in a dose-dependent manner (Figure 4a and b). Co-treatment with 10 μM compound **4** and 10 μM or 50 μM amitriptyline resulted in a significant reduction in the number of galectin-3 dots compared to amitriptyline treatment alone (Figure 4a and b), which strongly supports that compound **4** can act as an intracellular antagonist for galectin-3 in cell culture systems. The experimental concentration of **4** was selected to achieve a significant effect and possibly reflects a relatively slow cellular uptake and intracellular concentration increase of **4** sufficient to block intracellular galectin-3.

Pharmacological intervention in a bleomycin-induced lung fibrosis mouse model

Galectin-3 has been shown to promote both macrophage M2 polarization[5b] and myofibroblast activation,[6b] *i.e.* in two key pro-fibrotic cell types. In the case of macrophage M2 polarization, galectin-3 association with CD98 on the macrophage cell surface, presumably within lattices, was shown to be a plausible molecular mechanism for regulating M2 activation via phosphatidylinositol 3-kinase (PI3K) activation.[5b] Analogously, transforming growth factor-ß (TGF-ß) receptor II has been shown to bind galectin-3 on cell surfaces, which was suggested to be a critical molecular mechanism for inducing myofibroblast activation.[6b] Furthermore, *in vivo* an intratracheal single-dose of the galectin-3 antagonist **4** (10 μg per mouse, 500 μg/kg) was

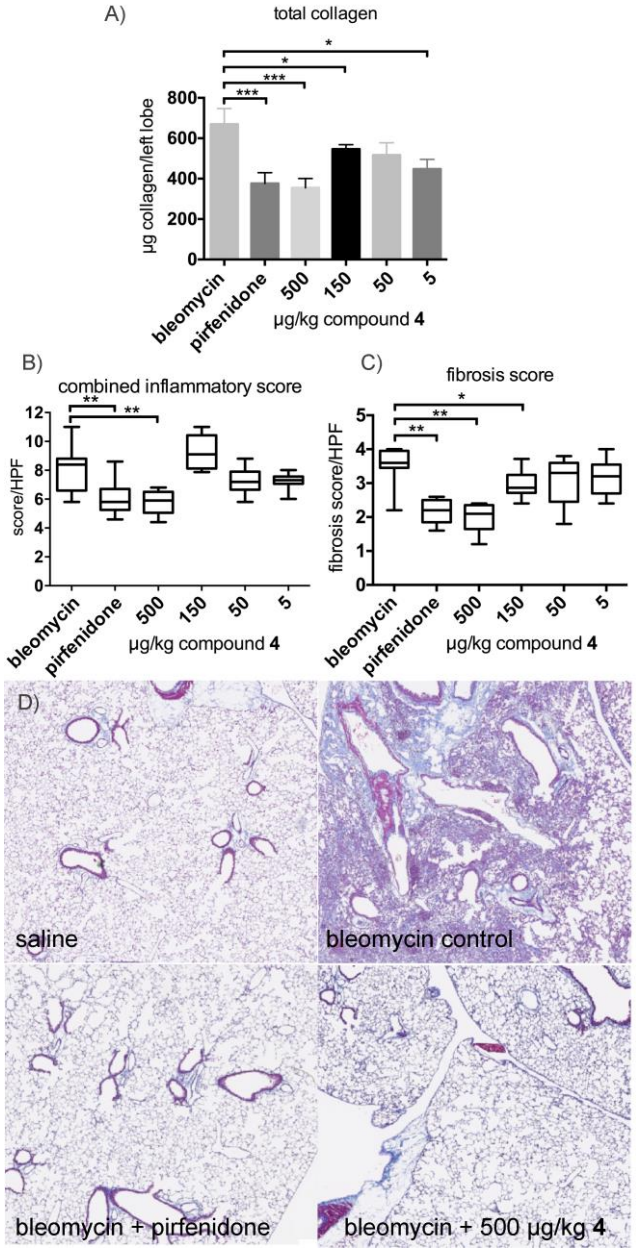


Figure 5. Effects of pirfenidone and **4** on bleomycin-induced lung fibrosis in mice. A) Total lung collagen, B) Histological inflammatory score, C) Histological fibrosis score. Results represent the mean ± SEM of n=8 mice per group. (* P<0.03, **P<0.05, ***P<0.01 statistically different from bleomycin control). D) Representative Masson's trichrome stained sections of mouse lung from uninjured saline control, bleomycin control and bleomycin treated with oral pirfenidone (200 mg/kg) or intratracheal **4** (500 µg/kg).

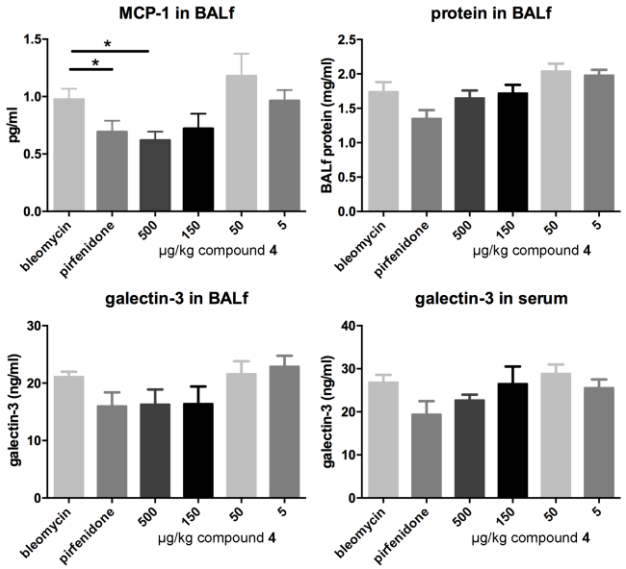


Figure 6. Effects of pirfenidone and compound **4** on BAL (bronchoalveolar lavage) fluid parameters in bleomycin-induced lung fibrosis in mice. Total protein measured by BCA reagent, MCP-1, and galectin-3 in BAL fluid and serum were measured by ELISA. Results represent the mean ± SEM of n=8 mice per group. (* P<0.05, statistically different from bleomycin control).

demonstrated to display an anti-fibrotic effect in a bleomycin-induced lung fibrosis mouse model.[6b] Hence, compound **4** can be hypothesized to possess dual anti-fibrotic effects by disrupting lattices with CD98 on M2 macrophages and with TGF-ß-RII on myofibroblasts and associated pro-fibrotic signaling. However, the single-dose experiment left questions unanswered concerning the *in vivo* dose-response efficacy of compound **4** and how this compared to alternative anti-fibrotic agents. Hence, we conducted a dose-response study of therapeutic administration of compound **4** in this model, in comparison with pirfenidone, one of only two recently approved drugs for treating idiopathic pulmonary fibrosis (IPF). Mice (n=8) received bleomycin sulphate (1.65 mg/kg intratracheally), resulting in inflammation and subsequent fibrosis development, followed by either 200 mg/kg pirfenidone twice daily orally from days 18-24 or compound **4** at 500, 150, 50 or 5 µg/kg intratracheally as a single administration every second day (days 18, 20, 22 and 24). Lung collagen content and histopathology was determined on day 26. Pirfenidone (200mg/kg) significantly reduced bleomycin-induced collagen accumulation from 670±77 to 375±53 µg collagen/lobe (P<0.01), as did 500 and 150 µg/kg of compound **4** (355±46 and 546±22 µg collagen/lobe P<0.01, P<0.03, respectively) (Figure 5). In addition, compound **4** at 500 and 150 µg/kg doses and pirfenidone significantly decreased the fibrosis score. Hence, when delivered directly into the lung, compound **4** achieves efficacy at much lower concentrations compared to orally delivered pirfenidone. The lower dose needed with administration of **4** could be due to the, not unexpectedly, improved lung targeting by intratracheal administration in combination with the high affinity of **4** for the target galectin-3 protein shown to be a key regulator of fibrosis

biochemistry. Lung availability of an orally administered compound is likely to be lower than that of an intratracheally administered compound, which may at least partly explain the need for a higher dose of oral pirfenidone to achieve the same efficacy as intratracheal **4**. In addition, compound **4** did not reduce protein in the BAL (bronchoalveolar lavage) fluid – an indication of vascular leakage – but both pirfenidone and compound **4** reduced MCP-1 (monocyte chemoattractant protein-1) levels (Figure 6). Compound **4** did not produce a clear significant decrease in galectin-3 levels in BAL fluid or serum. The absence of a significant decrease of galectin-3 in BAL fluid is likely due to the fact that BAL fluid samples are from whole lung and not only the fibrotic area. As non-fibrotic tissue has a background expression of galectin-3 this will influence the total galectin-3 levels in BAL fluid samples, therefore BAL fluid galectin-3 analysis may be underestimating the actual concentration of galectin-3 in the diseased areas of the lung.

Conclusions

Highly potent galectin-1 and 3 antagonists were discovered through synthesis optimization, and structural analysis of double C3 aryl-triazolyl-substituted thiodigalactosides. Low nM-affinities were reached for galectin-1 and 3 and some compounds displayed selectivity for individual galectins. Structural and mutational studies showed evidence that the exceptional affinity enhancement originated largely from the aryl-triazole moieties forming stacking interactions with protein π-systems (arginine side chains unpaired or ion-paired with glutamate or aspartate carboxylates) and in some cases fluorine-derived orthogonal multipolar interactions that endogenous glycoconjugate glycans do not form. The nature of the aryl-triazole moieties has a significant influence over galectin sub-type selectivities, which could also be explained by small, but significant, differences revealed in the structural studies. Overall, the results corroborate the promising strategy for discovery of high-affinity and selective lectin antagonists by exploring non-carbohydrate structural elements forming interactions that glycoside fragments of endogenous glycoconjugate ligands do not form with lectins. Hence, drug development targeting lectins may not necessarily involve a strategy of multimerizing ligands and antagonists to achieve sufficient affinities and the major challenges concerning pharmacokinetics, bioavailability, and toxicity/immunogenicity associated with multivalent antagonists may be avoided.

One antagonist (**4**) was evidenced to have intracellular availability and activity as it blocked amitriptyline-induced vesicle damage in breast carcinoma MCF-7 cells. While it still remains to be answered which glycoprotein binding partner is involved in the galectin-3 accumulation on damaged vesicles, the lysosome associated membrane proteins LAMP-1 and LAMP-2 may be candidates for this role as they have been shown to be galectin-3 ligands on the surface of tumor cells[24] and are thus possible candidate glycoprotein ligands in our model. Importantly, these observations suggest that intracellular galectin-3 glycoprotein-binding events occur and may be biologically relevant. Targeting such interactions with synthetic antagonists may be a viable strategy, although PK-ADMET properties obviously have to be improved for intracellular/systemic availability and therapeutic applications.

Furthermore, intratracheally delivered compound **4** attenuated bleomycin-induced lung fibrosis in a mouse model in a dose-dependent manner and possessed efficacy at significantly lower doses than the approved oral anti-fibrotic pirfenidone and thus compared favorably with pirfenidone. This may further support a dual molecular mechanistic hypothesis in which galectin-3-promoted macrophage and myofibroblast activation results in sustained pro-fibrotic cell signaling and scar formation.

Finally, five-membered aromatic heterocycles are common structural elements in many drugs and 1,2,3-triazoles in particular are readily synthesized, which render the compounds herein as promising leads for the development of novel galectin-targeting therapeutics that disrupt cellular signal-sustaining galectin-3 lattices as well as highly valuable tools for studying galectin biology and molecular mechanisms.

Experimental Section

Expression constructs, expression, and purification of recombinant galectins

Human galectin-1,[25] galectin-2,[26] galectin-3,[27] galectin-4N,[12a] galectin-4C,[12a] galectin-8N,[13b] and mouse galectin-7[28] were expressed and purified as earlier described. Human galectin-9N and galectin-9C were produced in *E. coli* BL21Star (DE3) cells (Invitrogen) and purified by affinity chromatography on lactosyl-Sepharose essentially as described for galectin-8.[13b] DNA encoding the genes of human galectin-9N and galectin-9C were cloned into the pET-32 Ek/LIC vector (Novagen, Madison, WI) according to the manufacturer's instructions. Briefly, I.M.A.G.E. clone 2208156 (ATCC) was used as template together with the following polymerase chain reaction (PCR) primers. The vector used for galectin-9N encoded the *N*-terminal 170 amino acids of galectin-9 and thioredoxin with the primers forward: 5`- GAC GAC GAC AAG ATG ATG GGT TCA GCG GTT CCC AGG-3´, forward 2: 5´- GAG GAG AAG CCC GGT TCA GGA AAC AGA CAG GCT GGG AGA ACGG C-3´, and reverse: 5´- GAG GAG AAG CCC GGT GCC GCC TAT GTC TGC ACA TGG G-3´. The vector used for galectin-9C encoded the *C*-terminal amino acids 205-355 of galectin-9 and thioredoxin with the primers forward: 5´-GAC GAC GAC AAG ATG GGA CAG ATG TTC TCT ACT CCC-3´ and reverse: 5´-GAG GAG AAG CCC GGT GCG GCC TAT GTC TGC ACA TGG G-3´. The bacteria were grown (37°C, 200 rpm) in LB (Luria-Bertani) medium with ampicillin (1mg/l) overnight, followed by induction with 1mM isopropyl thio-β-D-galactoside (IPTG) for 4 h (29°C, 200 rpm). The culture was centrifuged (15 min, 5000 rpm, 4°C) and the pellet was dissolved in 50 ml MEPBS (phosphate buffered saline with 2 mM EDTA and 4 mM β-mercaptoethanol) and sonicated 10-20 x 30 s on ice. The sonicated bacteria were centrifuged (30 min, 12000 rpm, 4°C) and the supernatant was submitted to affinity chromatography using a lactosyl-Sepharose column washed with MEPBS and a pre-elution of 7,5 mM lactose. The bound proteins were eluted with Lac-MEPBS (MEPBS with 150 mM lactose) as elution buffer. Removal of lactose was done by chromatography on a PD-10 column (Amersham Biosciences) and with repeated ultrafiltration using Centriprep (Amicon).

Competitive fluorescence polarization experiments determining galectin affinities

Fluorescence polarization experiments were performed on a POLARStar plate reader with software FLUOstar Galaxy software or a PheraStarFS plate reader with software PHERAstar Mars version 2.10 R3 (BMG, Offenburg, Germany) and fluorescence anisotropy of fluorescein tagged probes measured with excitation at 485 nm and emission at 520 nm. K_d values were determined in PBS as previously described[12, 29] with specific conditions for each galectin as described below. Compounds **3**-**10** were dissolved in neat DMSO at 100 mM and diluted in PBS to 3-6 different concentrations to be tested in duplicates. K_d average and SEM were calculated from 4 to 25 single point measurements showing between 30-70% inhibition.

Galectin-1 affinities: Experiments were done at 20°C with galectin-1 at 0.50 μM and the fluorescent probe 3,3'-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-3'-(3,5-dimethoxy-benzamido)-1,1'-sulfanediyl-di-β-D-galactopyranoside[25] at 0.10 μM.

Galectin-2 affinities: Experiments were done at 20°C with galectin-2 at 10 μM and the fluorescent probe 3,3'-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-3'-(3,5-dimethoxy-benzamido)-1,1'-sulfanediyl-di-β-D-galactopyranoside at 0.10 μM.

Galectin-3 affinities: Experiments were done at 20°C with galectin-3 at 0.20 μM and the fluorescent probe 3,3'-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-3'-(3,5-dimethoxybenzamido)-1,1'-sulfanediyl-di-β-D-galactopyranoside at 0.02 μM or with galectin-3 at 1.0 μM and 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1-2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside[12] at 0.10 μM.

Galectin-4N affinities: Experiments were done at 20°C with galectin-4N at 3.0 μM and the fluorescent probe 3,3'-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-3'-(3,5-dimethoxy-benzamido)-1,1'-sulfanediyl-di-β-D-galactopyranoside at 0.10 μM.

Galectin-4C affinities: Experiments were done at 20°C with galectin-4C at 0.50 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1-2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside at 0.1 μM.

Galectin-7 affinities: Experiments were done at 4°C with galectin-7 at 2.00 μM and the fluorescent probe β-D-galactopyranosyl(1—4)-2-acetamido-2-deoxy-β-D-glucopyranosyl(1—3)-β-D-galactopyranosyl(1—4)-(*N*1-fluorescein-5-yl-carbonylaminomethylcarbonyl)-β-D-glucopyranosylamine[30] at 0.1 μM.

Galectin-8N affinities: Experiments were done at 20°C with galectin-8N at 0.40 μM and the fluorescent probe 2-(fluorescein-5-yl-carbonylamino)ethyl β-D-galactopyranosyl(1—4)-2-acetamido-2-deoxy-β-D-glucopyranosyl(1—3)-β-D-galactopyranosyl(1—4)-β-D-glucopyranoside[13b] at 0.1 μM.

Galectin-9N affinities: Experiments were done at 20°C with galectin-9N at 1.0 μM and the fluorescent probe 2-(fluorescein-5-yl-carbonylamino)ethyl β-D-galactopyranosyl(1—4)-2-acetamido-2-deoxy-β-D-glucopyranosyl(1—3)-β-D-galactopyranosyl(1—4)-β-D-glucopyranoside at 0.1 μM.

Galectin-9C affinities: Experiments were done at 20°C with galectin-9C at 2.0 μM and the fluorescent probe 3,3'-dideoxy-3-[4-(fluorescein-5-yl-carbonylaminomethyl)-1*H*-1,2,3-triazol-1-yl]-3'-(3,5-dimethoxy-benzamido)-1,1'-sulfanediyl-di-β-D-galactopyranoside at 0.10 μM.

Galectin-3 R144S affinities: Experiments were done at 20°C with galectin-3 R144S at 0.30 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1–2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside at 0.02 μM.

Galectin-3 R144K affinities: Experiments were done at 20°C with galectin-3 R144K at 0.40 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1–2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside at 0.02 μM.

Galectin-3 R186S affinities: Experiments were done at 20°C with galectin-3 R186S at 3.50 μM and the fluorescent probe 2-(fluorescein-5/6-yl-carbonyl)-aminoethyl 2-acetamido-2-deoxy-α-D-galactopyranosyl-(1–3)-[α-L-fucopyranosyl-(1–2)]-β-D-galactopyranosyl-(1–4)-β-D-glucopyranoside at 0.1 μM.

Galectin-3 R186K affinities: Experiments were done at 20°C with galectin-3 R186K at 0.90 μM and the fluorescent probe β-D-galactopyranosyl(1—4)-2- acetamido-2-deoxy-β-D-glucopyranosyl(1—3)-β-D-galactopyranosyl(1—4)-(*N*1-fluorescein-5-yl-carbonylaminomethylcarbonyl)-β-D-glucopyranosylamine at 0.1 μM.

Crystallization

Compounds **4**, **5**, and **12** were prepared in the galectin-3 crystallization conditions by initially solubilizing in 55% w/v polyethylene glycol (PEG 6000), before addition of other crystallisation reagents to give a final concentration of 20 mM of **4**, **5**, and **12** in the galectin-3 crystallisation condition (31% w/v PEG 6000, 100 mM Tris-HC pH 7.5, 100 mM $MgCl_2$ for galectin-3). Galectin-3-CRD lactose or galactose co-crystals (prepared as previously described[31]) were soaked for 2–8 days in drop containing a 1:1 ratio of the ligand-containing crystallisation condition and 20 mg/mL human galectin-3-CRD in 10 mM Tris-HCl pH 7.5 (pre-equilibrated co-crystallisation drops that had not produced crystals).

X-ray diffraction analysis and structure determination

X-ray diffraction data sets were collected at room temperature from human galectin-3-CRD crystals mounted in 0.7 mm quartz capillaries on a ProteumR (Bruker AXS, Madison, WI, USA) diffractometer with a MacScience M06X^{CE} rotating-anode generator (wavelength 1.5418 Å) equipped with a SMART6000 CCD detector. X-ray diffraction data were integrated using SAINT (Bruker AXS, Madison, WI, USA) and scaled and merged using SCALA[32] within the CCP4 suite of crystallographic software.[33] Structures were solved by initial rigid body refinement using a previously published galectin-3-CRD structure (1A3K),[34] with ligand and waters removed, as the initial model. TLS and restrained refinement was performed using REFMAC5.[35] Anomalously scattering elements were identified using single-wavelength anomalous dispersion log-likelihood gradient maps (SAD LLG maps); calculated using Phaser [36] (in experimental phasing mode within CCP4) in the 'SAD with molecular replacement partial structure' mode with purely anomalous scatterers and zero LLG-map completion cycles using the current model and F+ and F− structure factor amplitudes as input. Visualization of electron density and model building was performed using Coot.[37] Ligand geometry topologies for refinement were initially created by REFMAC5 within CCP4 (LIBCHECK) or using the Dundee PRODRG2 Server.[38] In most cases minor to moderate manual editing of the automatically

generated topologies was performed to ensure correct atom and bond types. Model validation and analysis was performed using MolProbity.[39] Figures were created using the CCP4 molecular-graphics project (CCP4MG).[40]

Accession codes

PDB: The atomic coordinates and structure factors of galectin-3 in complex with **4**, **5**, and **12** have been deposited with accession codes 5E89, 5E8A, and 5E88, respectively.

Site-directed Mutagenesis

Mutants of human galectin-3 were made using the QuickChange® II site-directed mutagenesis kit (Stratagene, Amsterdam, The Netherlands), produced in *E.coli* BL21Star (DE3) cells (Invitrogen, Lidingö, Sweden) and purified by affinity chromatography on lactosyl-Sepharose as previously described.[41] Mutagenic primers for PCR were as follows: Gal-3R186K (AGA→AAA) sense (5´-CTG GGG AAG GGA AGA AAA ACA GTC GGT TTT CCC-3´) and antisense (5´-GGG AAA ACC GAC TGT TTT TCT TCC CTT CCC CAG-3´) and Gal-3R144K (AGA→AAA) sense (5´-GAA GCC CAA TGC AAA CAA AAT TGC TTT AGA TTT CCA AAG AG-3´) and antisense (5´-CTC TTT GGA AAT CTA AAG CAA TTT TGT TTG CAT TGG GCT TC-3´). Successful mutagenesis was confirmed by sequencing by GATC Biotech (Konstanz, Germany) in the forward direction from the T7 promotor primer and in the reverse direction from the pET-RP primer. Galectin-3 R144S and R186S were prepared as earlier reported.[41]

Cell culture and immunocytochemistry

MCF-7 cells were maintained in RPMI-1640 (Biochrom, Berlin, Germany) supplemented with 10% fetal bovine serum (Biochrom, Berlin, Germany), 10 µg/mL insulin (Sigma-Aldrich, Stockholm, Sweden), 100 µg/mL streptomycin and 100 units/mL penicillin (Hyclone). The cells were kept in a 37 °C humidified incubator supplied with 5% CO_2 in air. For the experiments stock solutions of **4** mM compound 4 in 40% dimethyl sulfoxide (DMSO) was used, while for amitriptyline (Sigma-Aldrich) stock solutions of 20 mM was made in sterile water. Both compound **4** and amitriptyline were serially diluted in RPMI-1640 before treatment of cells, such that the DMSO concentration did not exceed 0.1% v/v. MCF-7 (10^5 cells) were seeded onto sterile coverslips (placed in multiwell plates) and cultured for 24 hours. Cells were then treated with either 10 or 50 µM amitriptyline either alone or in combination with 10 µM of compound **4** for 24 hours. After fixation with 2% paraformaldehyde in phosphate-buffered saline (PBS) for 10 minutes, cells were permeabilized using 0.4% v/v Triton X-100 in PBS for 5 minutes. Non-specific binding was inhibited by blocking the cells with blocking buffer (1% w/v BSA, 0.1% v/v Tween 20 in PBS) for 10 minutes. Cells were then incubated with rat anti-mouse galectin-3 antibody (anti-Mac-2[42]) in a humidified chamber for 1 hour at room temperature. After three washes with PBS, goat anti-rat Alexa Fluor® 594 (Invitrogen, Carlsbad, USA) was added. Hoechst (10 ng/mL) was used to stain the nuclei. Cells were visualized by obtaining z-stacks of high magnification single optical planes using a LSM510 confocal laser scanning microscope (Carl Zeiss Microscopy GmbH, Oberkochen, Germany), conjugated with Hamamatsu R6357 (Hamamatsu Photonics K.K., Hamamatsu, Japan) photomultiplier. Galectin-3 dots were counted manually using ImageJ 1.47v and the plug-in Cell Counter (Wayne Rasband, National Institutes of Health, USA). Bar graphs representing galectin-3 dots/nuclei are expressed as mean values of different image areas ± SEM. For measuring statistical significance between a pair of data sets, Student's t-test (two tailed, unpaired) was employed. *P<0.05* was considered to be significant.

Bleomycin-induced fibrosis

Bleomycin was purchased from Apollo Scientific and reconstituted in sterile saline at a concentration of 0.66 mg/mL and aliquots were stored at -20°C. Pirfenidone was purchased from Tocris Biochemicals and was dissolved in 0.5% carboxymethyl cellulose (Sigma Aldrich) to a concentration of 20 mg/mL. Compound **4** was dissolved in 100% DMSO at a concentration of 10 mg/mL and aliquots stored at -20°C. For each day, compound **4** for instillation was diluted in sterile saline to give a final concentration of DMSO in the instillate of 2%. Female C57/Bl6 mice 10 weeks of age were purchased from Charles River and were maintained in 12-hour light/12-hour dark cycles with free access to food and water. All procedures were performed in accordance with Home Office guidelines (Animals (Scientific Procedures) Act 1986).

Mice were randomised into 8 treatment groups (n=8) and were anaesthetised with isofluorane and 33 µg bleomycin in 50 µL sterile saline was instilled into the lungs. A control group received 50 µL sterile saline. Mice were monitored closely over the next 26 days. Pirfenidone treated mice received pirfenidone 200 mg/kg by oral gavage twice daily from days 18-24. Mice treated with compound **4** received 50 µL intratracheally commencing day 18 every 48 hours for a total of 4 administrations. Control mice received vehicle (2% DMSO). Mice were culled on day 26. The lungs were perfused (via the right ventricle) with 5 ml saline and the lungs lavaged with 3 x 0.8 mL PBS containing 1 mM EDTA. BAL cells were combined and pelleted and lavage fluid from the first lavage was snap frozen. The lungs were removed and the entire left lobe removed and stored at -80°C for analysis of total collagen. Two upper right lobes were removed and snap frozen and stored at -80°C for subsequent RNA analysis. The remaining lung was inflated with 10% formalin and fixed for 24 hours prior to removal into 70% ethanol before embedding in paraffin wax for histological examination.

Total lung collagen

Frozen left lobes were thawed, weighed and minced finely with scissors and placed in 5ml of 3mg/ml pepsin in 0.5M acetic acid. Samples were incubated for 24 hours at 4°C and 0.2 mL of cleared extract was incubated with 0.8 mL Sircol reagent (Biocolor) for 60 minutes at room temperature. Collagen was sedimented by centrifugation at 13000 rpm for 5 minutes and the pellets resuspended in 0.5 mL of 0.5M NaOH. Samples were examined for absorbance at 560 nm with reference to a collagen standard curve.

Estimation of vascular leakage

Vascular leakage was determined by measuring total protein in the lavage fluid by BCA assay (Pierce) using bovine serum albumin as standard.

Histological lung inflammation and fibrosis score

Fibrosis and histological score was carried out in Masson's trichrome stained sections. Inflammation (peribronchiolar, perivascular, and alveolar wall thickness) scored in > 5 random fields at magnification X630 using the following system (peribronchiolar and perivascular, 1 = no cells, 2 = <20 cells, 3 = 20 – 100 cells, 4 = > 100 cells; alveolar wall thickness, 1 = no cells, 2 = 2 – 3 cells thick, 3 = 4 – 5 cells thick, 4 = > 5 cells thick). The combined inflammatory score is the sum of these scores. Fibrosis score was evaluated as the area of the section positively stained for collagen (1 = none, 2 = <10%, 3 = < 50%, 4 = > 50%). Only fields where the majority of the field is composed of alveoli were scored.

Determination of galectin-3 levels in BAL and serum by ELISA

Samples of BAL fluid and serum were assayed for galectin-3 and MCP-1 levels by ELISA (R&D systems)..

Acknowledgements

This work was supported by grants to U.J.N. and H.L. from the Swedish Research Council (Grants No. 621-2003-4265, 621-2006-3985, and 621-2009-5326), the programs “Glycoconjugates in Biological Systems” and “Chemistry for Life Sciences” sponsored by the Swedish Strategic Research Foundation, the foundation “Olle Engkvist Byggmästare”, the Royal Physiographic Society, Lund, by the European Community's Seventh Framework Program (FP7-2007-2013) under grant agreement n° HEALTH-F2-2011-256986–project acronym PANACREAS, and a project grant awarded by the Knut and Alice Wallenberg Foundation (KAW 2013.0022). The *in vivo* experiments were supported by Galecto Biotech AB, Sweden. H.B gratefully acknowledges the financial support from the Cancer Council Queensland, Australia (grants : ID1043716 and ID1080845). R.J.P. gratefully acknowledges the financial support from the Dutch Technology Foundation STW, applied science division of NOW. Barbro Kahl-Knutsson is acknowledged for performing fluorescence anisotropy experiments.

Keywords: galectin • antagonist • thiodigalactoside • vesicle • fibrosis

[1] P. Lajoie, J. G. Goetz, J. W. Dennis, I. R. Nabi, *J. Cell Biol.* **2009**, *185*, 381-385.

[2] C. Boscher, J. W. Dennis, I. R. Nabi, *Curr. Opin. Cell Biol.* **2011**, *23*, 383-392.

[3] F. T. Liu, G. A. Rabinovich, *Ann. N. Y. Acad. Sci.* **2010**, *1183*, 158-182.

[4] F.-T. Liu, G. A. Rabinovich, *Nat. Rev. Cancer* **2005**, *5*, 29-41.

[5] a) C.-I. Lin, E. E. Whang, D. B. Donner, X. Jiang, B. D. Price, A. M. Carothers, T. Delaine, H. Leffler, U. J. Nilsson, V. Nose, F. D. Moore, D. T. Ruan, *Mol. Cancer Res.* **2009**, *7*, 1655-1662; b) A. C. MacKinnon, S. L. Farnworth, P. S. Hodkinson, N. C. Henderson, K. M. Atkinson, H. Leffler, U. J. Nilsson, C. Haslett, S. J. Forbes, T. Sethi, *J. Immunol.* **2008**, *180*, 2650-2658.

[6] a) V. V. Glinsky, G. Kiriakova, O. V. Glinskii, V. V. Mossine, T. P. Mawhinney, J. R. Turk, A. B. Glinskii, V. H. Huxley, J. E. Price, G. V. Glinsky, *Neoplasia* **2009**, *11*, 901-909; b) A. C. Mackinnon, M. A. Gibbons, S. L. Farnworth, H. Leffler, U. J. Nilsson, T. Delaine, A. J. Simpson, S. J. Forbes, N. Hirani, J. Gauldie, T. Sethi, *Am. J. Resp. Crit. Care Med.* **2012**, *185*, 537-546.

[7] a) B. A. Salameh, H. Leffler, U. J. Nilsson, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3344-3346; b) B. A. Salameh, A. Sundin, H. Leffler, U. J. Nilsson, *Bioorg. Med. Chem.* **2006**, *14*, 1215-1220; c) P. Sörme, P. Arnoux, B. Kahl-Knutsson, H. Leffler, J. M. Rini, U. J. Nilsson, *J. Am. Chem. Soc.* **2005**, *127*, 1737-1743; d) P. Sörme, Y. Qian, P.-G. Nyholm, H. Leffler, U. J. Nilsson, *ChemBioChem* **2002**, *3*, 183-189; e) J. Tejler, B. Salameh, H. Leffler, U. J. Nilsson, *Org. Biomol. Chem.* **2009**, *7*, 3982-3990; f) S. Fort, H. S. Kim, O. Hindsgaul, *J. Org. Chem.* **2006**, *71*, 7146-7154; g) D. Giguère, S. André, M.-A. Bonin, M.-A. Bellefleur, A. Provencal, P. Cloutier, B. Pucci, R. Roy, H.-J. Gabius, *Bioorg. Med. Chem.* **2011**, *19*, 3280-3287; h) D. Giguere, M. Bonin, P. Cloutier, R. Patnam, C. Stpierre, S. Sato, R. Roy, *Bioorg. Med. Chem.* **2008**, *16*, 7811-7823; i) D. Giguere, R. Patnam, M.-A. Bellefleur, C. St.-Pierre, S. Sato, R. Roy, *Chem. Commun.* **2006**, 2379-2381; j) D. Giguere, S. Sato, C. Stpierre, S. Sirois, R. Roy, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1668-1672; k) M. F. Marchiori, D. E. P. Souto, L. O. Bortot, J. F. Pereira, L. T. Kubota, R. D. Cummings, M. Dias-Baruffi, I. Carvalho, V. L. Campo, *Bioorg. Med. Chem.* **2015**, *23*, 3414-3425; l) G. A. Rabinovich, A. Cumashi, G. A. Bianco, D. Ciavardelli, I. Iurisci, M. D'Egidio, E. Piccolo, N. Tinari, N. Nifantiev, S. Iacobelli, *Glycobiology* **2005**, *16*, 210-220; m) S. R. Rauthu, T. C. Shiao, S. André, M. C. Miller, É. Madej, K. H. Mayo, H.-J. Gabius, R. Roy, *ChemBioChem* **2015**, *16*, 126-139.

[8] a) I. Cumpstey, E. Salomonsson, A. Sundin, H. Leffler, U. J. Nilsson, *Chem. Eur. J.* **2008**, *14*, 4233-4245; b) I. Cumpstey, A. Sundin, H. Leffler, U. J. Nilsson, *Angew. Chem. Int. Ed.* **2005**, *44*, 5110-5112; c) T. Delaine, I. Cumpstey, L. Ingrassia, M. Le Mercier, P. Okechukwu, H. Leffler, R. Kiss, U. J. Nilsson, *J. Med. Chem.* **2008**, *51*, 8109-8114; d) B. A. Salameh, I. Cumpstey, A. Sundin, H. Leffler, U. J. Nilsson, *Bioorg. Med. Chem.* **2010**, *18*, 5367-5378.

[9] M. Van Scherpenzeel, E. E. Moret, L. Ballell, R. M. J. Liskamp, U. J. Nilsson, H. Leffler, R. J. Pieters, *ChemBioChem* **2009**, *10*, 1724-1733.

[10] H. van Hattum, H. M. Branderhorst, E. E. Moret, U. J. Nilsson, H. Leffler, R. J. Pieters, *J. Med. Chem.* **2013**, *56*, 1350-1354.

[11] T. L. Lowary, O. Hindsgaul, *Carbohydr. Res.* **1994**, *251*, 33-67.

[12] a) P. Sörme, B. Kahl-Knutsson, M. Huflejt, U. J. Nilsson, H. Leffler, *Anal. Biochem.* **2004**, *334*, 36-47; b) P. Sörme, B. Kahl-Knutsson, U. Wellmar, U. J. Nilsson, H. Leffler, *Meth. Enzymol.* **2003**, *362*, 504-512.

[13] a) S. Carlsson, M. C. Carlsson, H. Leffler, *Glycobiology* **2007**, *17*, 906-912; b) S. Carlsson, C. T. Öberg, M. C. Carlsson, A. Sundin, U. J. Nilsson, D. Smith, R. D. Cummings, J. Almkvist, A. Karlsson, H. Leffler, *Glycobiology* **2007**, *17*, 663-676.

[14] K. Bum-Erdene, I. A. Gagarinov, P. M. Collins, M. Winger, A. G. Pearson, J. C. Wilson, H. Leffler, U. J. Nilsson, I. D. Grice, H. Blanchard, *Chembiochem* **2013**, *14*, 1331-1342.

[15] I. Cumpstey, E. Salomonsson, A. Sundin, H. Leffler, U. J. Nilsson, *ChemBioChem* **2007**, *8*, 1389-1398.

[16] a) K. Müller, C. Faeh, F. Diederich, *Science* **2007**, *317*, 1881-1886; b) R. Paulini, K. Müller, F. Diederich, *Angew. Chem. Int. Ed.* **2005**, *44*, 1788-1805; c) M. Zürcher, F. Diederich, *J. Org. Chem.* **2008**, *73*, 4345-4361.

[17] P. Zhou, J. Zou, F. Tian, Z. Shang, *J. Chem. Inf. Mod.* **2009**, *49*, 2344-2355.

[18] J. C. Biffinger, H. W. Kim, S. G. DiMagno, *ChemBioChem* **2004**, *5*, 622-627.

[19] D. N. Cooper, *Biochim Biophys Acta* **2002**, *1572*, 209-231.

[20] a) S. Aits, J. Kricker, B. Liu, A.-M. Ellegaard, S. Hämälistö, S. Tvingsholm, E. Corcelle-Termeau, S. Høgh, T. Farkas, A. H. Jonassen, I. Gromova, M. Mortensen, M. Jäättelä, *Autophagy* **2015**, *11*, 1408-1424; b) I. Paz, M. Sachse, N. Dupont, J. Mounier, C. Cederfur, J. Enninga, H. Leffler, F. Poirier, M. C. Prevost, F. Lafont, P. Sansonetti, *Cell Microbiol.* **2010**, *12*, 530-544.

[21] T. L. Thurston, M. P. Wandel, N. von Muhlinen, A. Foeglein, F. Randow, *Nature* **2012**, *482*, 414-418.

[22] N. H. Petersen, O. D. Olsen, L. Groth-Pedersen, A. M. Ellegaard, M. Bilgin, S. Redmer, M. S. Ostenfeld, D. Ulanet, T. H. Dovmark, A. Lonborg, S. D. Vindelov, D. Hanahan, C. Arenz, C. S. Ejsing, T. Kirkegaard, M. Rohde, J. Nylandsted, M. Jaattela, *Cancer Cell* **2013**, *24*, 379-393.

[23] S. Aits, J. Kricker, B. Liu, A. M. Ellegaard, S. Hamalisto, S. Tvingsholm, E. Corcelle-Termeau, S. Hogh, T. Farkas, A. H. Jonassen, I. Gromova, M. Mortensen, M. Jaattela, *Autophagy* **2015**, 0.

[24] V. Sarafian, M. Jadot, J. M. Foidart, J. J. Letesson, F. van den Brûle, V. Castronovo, R. Wattiaux, S. Wattiaux-De Coninck, *Int. J. Cancer* **1998**, *75*, 105-111.

[25] E. Salomonsson, A. Larumbe, J. Tejler, E. Tullberg, H. Rydberg, A. Sundin, T. Khabut, T. Frejd, Y. D. Lobsanov, J. M. Rini, U. J. Nilsson, H. Leffler, *Biochemistry* **2010**, *49*, 9518-9532.

[26] M. A. Gitt, S. M. Massa, H. Leffler, S. H. Barondes, *J. Biol. Chem.* **1992**, *267*, 10601-10606.

[27] S. M. Massa, D. N. Cooper, H. Leffler, S. H. Barondes, *Biochemistry* **1993**, *32*, 260-267.

[28] T. Magnaldo, D. Fowlis, M. Darmon, *Differentiation* **1998**, *63*, 159-168.

[29] I. Cumpstey, S. Carlsson, H. Leffler, U. J. Nilsson, *Org. Biomol. Chem.* **2005**, *3*, 1922-1932.

[30] C. T. Öberg, S. Carlsson, E. Fillion, H. Leffler, U. J. Nilsson, *Bioconj. Chem.* **2003**, *14*, 1289-1297.

[31] P. M. Collins, K. I. P. J. Hidari, H. Blanchard, *Acta Crystallogr. D Biol. Crystallogr.* **2007**, *63*, 415-419.

[32] P. Evans, *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 72-82.

[33] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235-242.

[34] J. Seetharaman, A. Kanigsberg, R. Slaaby, H. Leffler, S. H. Barondes, J. M. Rini, *J. Biol. Chem.* **1998**, *273*, 13047-13052.

[35] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355-367.

[36] R. J. Read, A. J. McCoy, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 338-344.

[37] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486-501.

[38] A. W. Schuttelkopf, D. M. van Aalten, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1355-1363.

[39] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12-21.

[40] S. McNicholas, E. Potterton, K. S. Wilson, M. E. Noble, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 386-394.

[41] E. Salomonsson, M. Carlsson, V. Osla, R. Hendus-Altenburger, B. Kahl Knutson, C. Oberg, A. Sundin, R. Nilsson, E. Nordberg-Karlsson, U. Nilsson, A. Karlsson, J. M. Rini, H. Leffler, *J. Biol. Chem.* **2010**, *285*, 35079-35091.

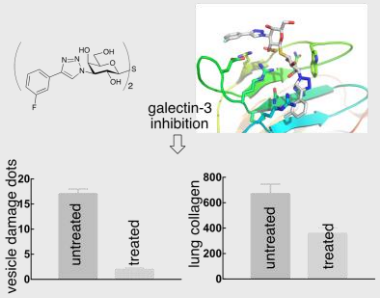
[42] M. K. Ho, T. A. Springer, *J. Immunol.* **1982**, *128*, 1221-1228.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

FULL PAPER

Dot removal: Synthetic ligands capitalizing on unnatural lectin interactions efficiently inhibit intracellular galectin-3 accumulation in dots on damaged vesicles and attenuates experimental lung fibrosis



((Insert TOC Graphic here: max. width: 5.5 cm; max. height: 5.0 cm))

T. Delaine, P. Collins, A. MacKinnon, G. Sharma, J. Stegmayr, V. K. Rajput, S. Mandal, I. Cumpstey, A. Larumbe, B. A. Salameh, B. Kahl-Knutsson, H. van Hattum, M. v. Scherpenzeel, R. J. Pieters, T. Sethi, H. Schambye, S. Oredsson, H. Leffler, H. Blanchard, and U. J. Nilsson**

Page No. – Page No.

Galectin-3-binding glycomimetics that strongly reduce bleomycin-induced lung fibrosis and modulate intracellular glycan recognition